Journal of Medical Virology 60:159–165 (2000)

High Prevalence of GB Virus C/Hepatitis G Virus in Kinshasa, Democratic Republic of Congo: A Phylogenetic Analysis

Hsin-Fu Liu,[1] Jean-Jacques Muyembe-Tamfum,[2] Karin Dahan,[3] Jan Desmyter,[4] and Patrick Goubau[1]*

[1]*Unit of Virology, Cliniques Universitaires St-Luc, Université Catholique de Louvain, Brussels, Belgium*
[2]*Department of Microbiology, Université de Kinshasa, Kinshasa, Congo*
[3]*Genetics, Cliniques Universitaires St-Luc, Université Catholique de Louvain, Brussels, Belgium*
[4]*Department of Microbiology and Immunology, Rega Institute for Medical Research, Katholieke Universiteit Leuven, Leuven, Belgium*

A prevalence of 10.3% of GB virus C (GBV-C)/hepatitis G virus (HGV) carriers was found in 97 pregnant women from Kinshasa, Congo (formerly Zaire), while prevalences of 1%, 4.1%, and 0% were found for hepatitis C virus, human immunodeficiency virus, and human T-lymphotropic virus respectively. Phylogenetic analysis of the ten GBV-C/HGV positives based on the 5′ non-coding region using three different methods identified consistently three GBV-C/HGV genotypes. Four main clades were found within the type 1 sequences. All the Congolese isolates are GBV-C/HGV type 1 in two different clades. The clustering of seven Congolese isolates was inconsistent in different methods. Further likelihood-mapping analysis showed a well-resolved phylogeny, confirming the clustering of the seven Congolese isolates with a Belgian strain representing a new clade in the GBV-C/HGV type 1 sequences. ***J. Med Virol. 60:159–165, 2000.***

KEY WORDS: GBV-C; HGV, 5′non-coding region; genotyping

INTRODUCTION

GB virus C (GBV-C) and hepatitis G virus (HGV) were identified independently from patients with viral hepatitis by two research groups [Simons et al., 1995; Linnen et al., 1996]. Sequence comparison of GBV-C and HGV showed that they have up to 95% amino acid and approximately 86% nucleotide sequence similarity, they are thus different isolates of the same virus [Zuckerman, 1996]. Although the overall genomic structure of GBV-C/HGV is similar to that of hepatitis C virus (HCV) which can cause chronic hepatitis, cirrhosis, and possibly hepatocellular carcinoma [Choo et al., 1989; Saito et al., 1990], the clinical significance of GBV-C/HGV infection and its pathogenic role in hepatitis or any other disease remains unclear. In the populations where both viruses are prevalent, double infections with HCV and GBV-C/HGV were frequently observed suggesting common transmission pathways. In some parts of Africa, a very high prevalence of HCV has been found [Kowo et al., 1995]. Therefore, an investigation of both viruses was undertaken among women attending a prenatal clinic in Kinshasa, the Democratic Republic of the Congo (former Zaire). Antibodies to human immunodeficiency virus (HIV) and human T-lymphotropic virus (HTLV) were also tested to compare with previous results in similar populations (Goubau P, unpublished data).

GBV-C/HGV is an enveloped positive-strand RNA virus with a genome size of about 9.3 kilobases belonging to the *Flaviviridae* family [Leary et al., 1996]. Phylogenetic analysis based on the 5′ non-coding region (5′NCR) sequence has identified at least 3 major GBV-C/HGV genotypes [Fukushi et al., 1996; Muerhoff et al., 1996]. In general, type 1 is more prevalent in Africa, type 2 in Europe and North America, and type 3 in Asia, but this geographical grouping is not absolute. Type 2 contains also strains from the other parts of the world such as Pakistan, Japan, and East Africa [Muerhoff et al., 1997; Smith et al., 1997]. Although type 2 can be divided into further subtypes 2a and 2b, the subtyping of type 1 is less certain [Liu et al., 1998]. In addition to the known 3 genotypes, a new fourth genotype which was found in Myanmar and Vietnam has also been reported recently [Naito et al., 1999].

Grant sponsor: Belgian Fund for Medical Scientific Research (FRSM); Grant number: 3.4523.98.

*Correspondence to: Patrick Goubau, Unit of Virology, Cliniques Universitaires St-Luc, Université Catholique de Louvain, Clos Chapelle-aux-Champs 30, B-1200 Brussels, Belgium.
E-mail: goubau@mblg.ucl.ac.be

Accepted 2 July 1999

In order to obtain a better understanding of the phylogenetic relationship among GBV-C/HGV type 1 isolates, analyzing more strains from Africa should be necessary. In this study, 10 nucleotide sequences of the 5′NCR of the viral genome were determined in carriers in Congo and phylogenetic analysis was carried out by different methods. The difference in clustering of GBV-C/HGV type 1 isolates given by different phylogenetic analysis methods is described.

MATERIALS AND METHODS

Serum Samples and Viral Screening

In 1996, 97 pregnant Congolese women attending a prenatal clinic in Kinshasa agreed to give blood for testing for blood-borne viruses. The samples were tested for the presence of GBV-C/HGV RNA and antibodies for HCV, HIV, and HTLV. HTLV was examined by passive particle agglutination (Serodia HTLV, Fujirebio, Tokyo, Japan) and reactive samples were confirmed by western blot (HTLV blot 2.4, Genelabs, Singapore). HIV screening was performed with a microparticle enzyme immunoassay (Axsym, Abbott Laboratories, North Chicago, IL) and confirmation was done by western blot (HIV blot 2.2, Genelabs). Antibodies to HCV were detected by a third generation enzyme linked immunosorbent assay (ELISA) (HCV 3.0 ELISA test system, Ortho Diagnostic Systems, Neckargemünd, Germany) and samples reacting by this first test were retested with a second third generation ELISA (HCV EIA 3.0, Abbott Laboratories). Samples are considered as positive if they react with both ELISAs. GBV-C/HGV RNA was detected by reverse transcription-polymerase chain reaction (RT-PCR) amplifying 2 different genomic regions (5′NCR and NS5) as described previously [Cornu et al., 1997]. GBV-C/HGV positive (with two primer sets) and indeterminate (positive with one primer set) samples were taken for further study. A questionnaire was completed for each woman concerning age, region of origin, parity, previous hospitalization, a history of transfusions, sexually transmitted diseases and present diseases.

RT-PCR and Sequencing

RNA was extracted from 50 μl of serum using the guanidinium isothiocyanate-phenol-chloroform method and used for cDNA synthesis as described previously [Cornu et al., 1997]. Five microliter of the cDNA was used in a nested polymerase chain reaction (PCR) amplifying the 5′NCR to part of the putative E1 gene of GBV-C/HGV genome. PCR products were purified from agarose gel and subjected to direct sequencing. PCR primers, reactive conditions, DNA purification, and sequencing method have been described elsewhere [Liu et al., 1998].

Phylogenetic Analysis

A basic local alignment search tool (BLAST) search was carried out to obtain the other GBV-C sequences in the same amplified region available in the EMBL/GenBank database. The sequences retrieved are shown

TABLE I. Origin of the GBV-C Isolates in the Phylogenetic (In Alphabetical Order)

Strain	Geographical origin	Accession number
AM	Belgium	Y15257
DH	Belgium	Y15255
DJ	Belgium	Y15256
DR	Belgium	Y15259
GBV-C (EA)	East Africa	U63715
GBV-C1	U.S.A.	U59518
GBV-C2	U.S.A.	U59519
GBV-C3	Greece	U59520
GBV-C4	Italy	U59521
GBV-C6	U.S.A.	U59523
GBV-C7	U.S.A.	U59524
GBV-C8	U.S.A.	U59525
GBV-C9	U.S.A.	U59526
GBV-C10	U.S.A.	U59527
GBV-C11	U.S.A.	U59528
GBV-C12	Greece	U59529
GBV-C13	Greece	U59530
GBV-C14	U.S.A.	U59531
GBV-C15	U.S.A.	U59532
GBV-C16	Europe	U59533
GBV-C17	Greece	U59534
GBV-C18	U.S.A.	U59535
GBV-C20	U.S.A.	U59537
GBV-C22	Japan	U59539
GBV-C23	Ghana	U59540
GBV-C24	Ghana	U59541
GBV-C25	Ghana	U59542
GBV-C26	Ghana	U59543
GBV-C27	Ghana	U59544
GBV-C28	Ghana	U59545
GBV-C29	Ghana	U59546
GBV-C31	Ghana	U59547
GBV-C32	Ghana	U59548
GBV-C33	Ghana	U59549
GBV-C34	Ghana	U59550
GBV-C35	Ghana	U59551
GBV-C36	Ghana	U59552
GBV-C37	Ghana	U59553
GBV-C38	Ghana	U59554
GBV-C39	Ghana	U59555
GBV-C40	Ghana	U59556
GBV-C41	Ghana	U59557
GBV-C42	Ghana	U59558
GP	Belgium	Y15261
GS185	Japan	D87262
GT110	Japan	D90600
GT230	Japan	D90601
HGV-IM71	Japan	AB008342
HGV-lw	Japan	D87255
MV	Belgium	Y15258
PA	Belgium	Y15260
PNF2161	U.S.A.	U44402
RM	Belgium	Y15262
TJ	Belgium	Y15263
VC	Belgium	Y15264
WA	Belgium	Y15265
WY	Belgium	Y15266
Z11	DR Congo	Y16435
Z19	DR Congo	Y16436
Z42	DR Congo	Y16437
Z53	DR Congo	Y16438
Z59	DR Congo	Y16439
Z60	DR Congo	Y16440
Z77	DR Congo	Y16441
Z78	DR Congo	Y16442
Z80	DR Congo	Y16443
Z91	DR Congo	Y16444

TABLE II. Sequence Divergence (in %) of 5′ Non-Coding Region of GB Virus C/Hepatitis G Virus Isolates From Congo

	Z11	Z19	Z42	Z53	Z59	Z60	Z77	Z78	Z80	Z91
Z11	0									
Z19	2.9	0								
Z42	6.1	5.4	0							
Z53	2.9	2.5	6.1	0						
Z59	7.4	6.8	5.1	7.5	0					
Z60	4.1	3.6	7.5	2.7	8.0	0				
Z77	2.7	2.5	6.3	1.2	7.6	2.5	0			
Z78	3.4	3.6	6.8	2.5	7.8	3.9	2.0	0		
Z80	4.1	4.6	7.8	4.1	8.3	4.6	3.9	3.7	0	
Z91	6.4	5.6	3.7	5.8	3.2	6.4	5.9	6.4	7.1	0

in Table I together with sequences from a previous study. A total of 67 sequences were aligned with the GeneWorks software (version 2.4, Intelligenetics, Geel, Belgium) followed by minimal manual editing. Phylogeny construction and evaluation were performed using the Phylip software package (version 3.572, University of Washington, Seattle, WA) [Felsenstein, 1989], with the neighbour-joining method (NJ), the Fitch and Wagner parsimony method (pars), and the maximum likelihood method (ML). An empirical transition/transversion ratio of 2.5 was estimated by the PUZZEL software (version 4.0, Universität München, Munich, Germany) [Strimmer and von Haeseler, 1996], and was used with the Felsenstein model to calculate the evolutionary distances. The robustness of the NJ and pars trees were statistically evaluated by bootstrap analysis with 1000 bootstrap samples [Felsenstein, 1985]. Since the ML method is already a statistical method (with a statistical evaluation of the branch length), no bootstrapping was done for it.

Likelihood-Mapping Analysis

Controversial phylogenies were further analyzed by likelihood-mapping analysis implemented in the PUZZEL software [Strimmer and von Haeseler, 1997]. The method is based on an analysis of the maximum likelihoods for the three fully resolved tree topologies that can be computed for four sequences. The three likelihoods are represented as points inside an equilateral triangle. The triangle is partitioned into different regions. The center of the triangle represents a star-like evolution whereas the three corners represent well-resolved phylogeny and the three intermediate regions between the corners reflect the difficulty in distinguishing between two of the three trees. For more than four sequences, the different strains can be grouped into four different subsets, for example, according to their position in a phylogenetic tree, and all possible quartets generated by drawing one sequence from each subset can be evaluated. The likelihoods for each subset of four sequences are mapped onto the triangle. The more points distribute in a certain region of a particular corner, the bigger the support for the tree topology joining the four subsets represented by that corner. On the other hand, if most points locate in the center of the triangle, the four subsets are independent and related by a star-like tree.

RESULTS

Prevalences of GBV-C/HGV, HCV, HIV, and HTLV

Ninety-seven Congolese pregnant women attending a prenatal clinic in Kinshasa were tested for the presence of GBV-C/HGV RNA by two sets of RT-PCRs amplifying the 5′NCR and the NS5a regions of the viral genome. The women came from different regions of Congo. Eight (Z11, Z19, Z42, Z59, Z60, Z77, Z78, Z91) were found to be positive in both PCRs while two (Z53, Z80) were only positive in the NS5a amplification. The GBV-C/HGV RNA prevalence was found to be 10.3% after the successful amplification of these two indeterminates in a subsequent nested PCR. The distribution of geographical regions of origin of the GBV-C/HGV positives was not different from the GBV-C/HGV negatives. There was no difference in transfusion history or recall of sexually transmitted diseases between the two groups. For HCV, only one (1%) was found to be positive who was also co-infected with GBV-C/HGV. Four of the 97 (4.1%) were HIV positive while they were all HTLV negative.

PCR and Sequence Comparison

The ten GBV-C/HGV positive and indeterminate samples were then taken in a nested PCR amplifying the 5′NCR to part of the putative E1 gene of the viral genome. The PCR fragments with the expected size of about 630 bp were amplified successfully and sequenced from all the 10 samples. The sequences were deposited in the EMBL database and were assigned the accession numbers Y16435 to Y16444. The overall nucleotide sequence variability among the 10 isolates ranged from 1.2% to 8.3% (Table II). The isolates Z53 and Z77 are the closest strains with a sequence similarity of 98.8% whereas the isolates Z59 and Z80 are the most divergent from one another with a sequence similarity of 91.7%.

Phylogenetic Analysis of the 5′NCR Sequences

A total of 67 GBV-C/HGV sequences were included in the phylogenetic analysis based on the 592 nucleotides 5′NCR fragment. For the ML method, only 63 sequences were analysed due to the computational capacity. A consistent tree topology was obtained with all methods and three GBV-C/HGV genotypes were ob-

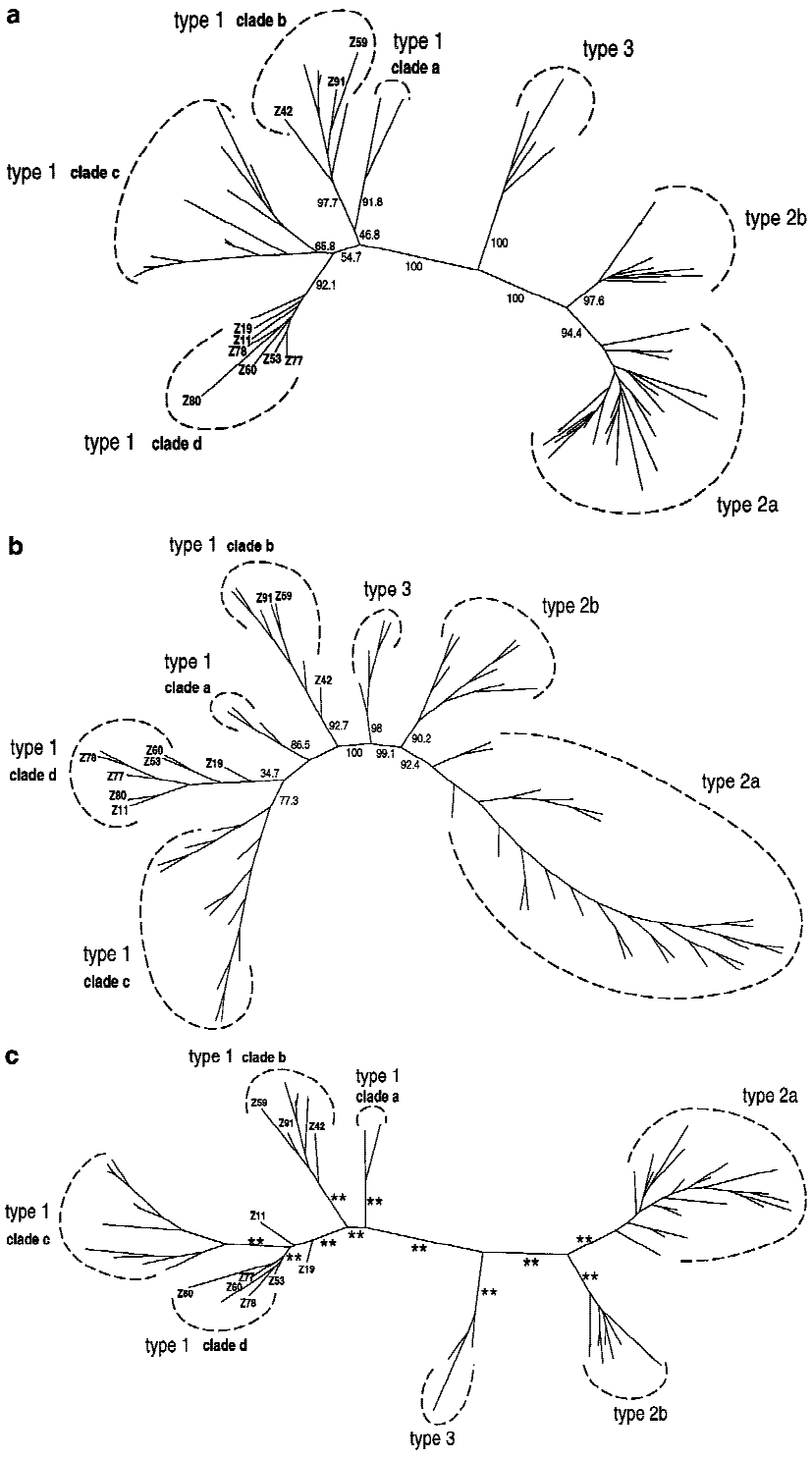

Fig. 1. Phylogenetic trees of the ten Congolese GB virus C/hepatitis G virus (GBV-C/HGV) isolates based on a 592 nucleotides fragment starting from the 5′ non-coding region to part of the E1 gene of the viral genome. Trees were constructed by **(a)** neighbour-joining method, **(b)** the Fitch and Wagner parsimony method, and **(c)** the maximum likelihood method. The bootstrap values (1000 bootstrap samples) are indicated beside the branches in percentage. In the maximum likelihood tree, double stars (**) represent a significantly positive branch (P<0.01). Only the names of the Congolese strains are shown.

served (Fig. 1). In general, types 1, 2, and 3 mainly consist of isolates from Africa, Europe or North America, and Asia respectively. Two monophyletic clades were found within the type 2 sequences and four main clades were observed within the type 1 sequences. The appearance of two subtypes (a and b) in the type 2 was well supported by all methods, whereas the clustering of the four type 1 clades was inconsis-

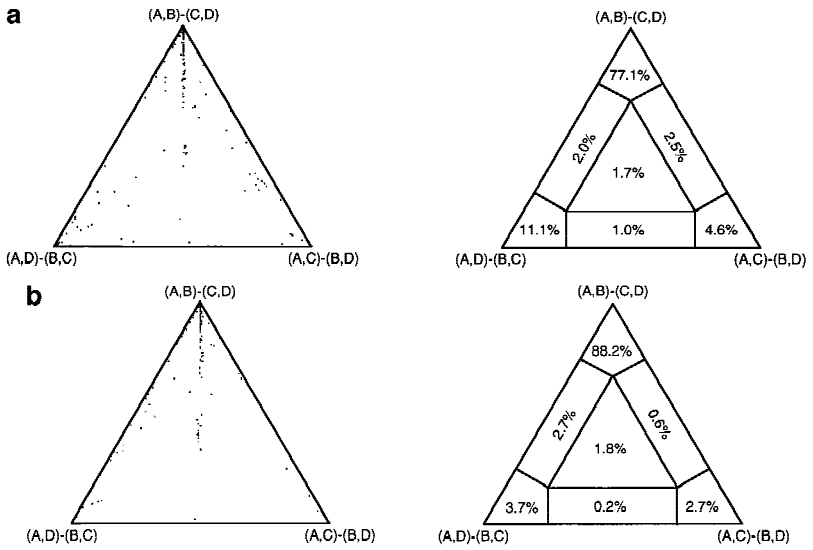

Fig. 2. Likelihood-mapping statistics for the clustering of **(a)** Z19 with GBV-C/HGV type 1 clade d and **(b)** Z 11 with clade d. The three likelihoods are represented as points inside the first triangle and the number of quartets in the seven areas are given in the second triangle. The assignments of the analysed sequences in the groups A, B, C, and D are given in the results section.

tent. In the NJ tree, clades a and b clustered together with bootstrap values of 46.8% (Fig. 1a). However, they fell into two independent branches in the pars (Fig. 1b) and the ML (Fig. 1c) trees. All the ten Congolese isolates are GBV-C/HGV type 1. Three (Z42, Z59, Z91) clustered in clade b together with four strains from Ghana in all methods with high bootstrap values (97.7% for the NJ, 92.7% for the pars, and $P<0.01$ for ML), whereas the clustering of the other seven Congolese isolates was inconsistent by different methods. In the NJ tree, these seven strains clustered together with an isolate from Belgium and formed a well supported monophyletic group (clade d) with a bootstrap value of 92.1%, while in the pars tree, clade d only obtained 34.7% bootstrap support. Moreover, two strains (Z11, Z19) were excluded from this branch in the ML tree. In spite of this, the branch length of clade d remained statistical significant ($P<0.01$) in the ML tree. The clustering of these Congolese isolates was not specific for the geographical origins of the women.

Likelihood-Mapping Analysis

Likelihood-mapping analyses can be used as a complementary approach to solve controversial phylogenies. For testing the inconsistent topologies of clade d, two sets of likelihood-mapping analyses were done. The first set assigned Z19 as group A, the rest of sequences in clade d as group B, the other sequences in clade c as group C, and all the sequences in clades a and b as group D. The second set has the same grouping assignment as in the first set, except that Z11 was assigned as group A and Z19 was assigned in group B together with the rest of sequences in clade d. The first set of likelihood-mapping statistics showed that 77.1% of the quartets were mapped in the top corner of the triangle (Fig. 2a) supporting the grouping of Z19 with clade d. The second set showed that 88.2% of the quartets were mapped in the top corner of the triangle (Fig. 2b) supporting the grouping of Z11 with clade d.

DISCUSSION

This study showed a stable picture for the prevalence of HIV and HTLV in this population over the past ten years, when compared to previous unpublished results, with 4.1% (95% C.I.: 1.13–10.22) for HIV and 0% (95% C.I.: 0.0-3.73) for HTLV (Goubau P, unpublished data). In Congo, HCV (1%) seems less prevalent than GBV-C/HGV (10.3%) based on the current results and also compared to the previously reported HCV prevalence of 6.4% by others [Tibbs et al., 1991]. Interestingly, the HCV positive woman in this study was also positive for GBV-C/HGV. The observed rate of 10.3% GBV-C/HGV RNA positivity is comparable to that observed in high-risk groups (e.g. in hemodialysis patients) in industrialized countries [Masuko et al., 1996; Cornu et al., 1997]. A very similar prevalence of GBV-C/HGV RNA (10.4%) has also been found in a rural South African population [Tucker et al., 1997]. It may be concluded that GBV-C/HGV is much more readily transmitted than HCV. The true GBV-C/HGV prevalence should be higher than the present one since only viral RNA was tested, not antibody.

Three different types of phylogenetic analysis were carried out using the NJ, the pars, and the ML methods to increase the reliability of the derived topologies. Since the algorithms for constructing the phylogenetic trees of these three methods are based on different evolutionary assumptions, a consistent topology with all three methods increases its reliability. Our analysis was based on the 5′NCR region to part of the E1 gene of the viral genome because this region is more informative for phylogeny reconstruction. Phylogenetic

analysis of these GBV-C/HGV sequences consistently identified three GBV-C/HGV genotypes. The subtyping of 2a and 2b was well supported by all methods and was concordant with the results of others [Muerhoff et al., 1997; Smith et al., 1997]. In contrast, the clustering of the four clades in type 1 is more problematical. The branching patterns were inconsistent in different methods. Previously a possible subtyping of GBV-C/HGV type 1 into 1a and 1b was obtained by means of the ML method with a statistically significant support [Liu et al., 1998]. Smith and collaborators [1997], on the basis of shorter sequences, also suggested that subgrouping might exist among genotype 1 strains. In the present study, addition of the Congolese strains led to the previously observed 1a cluster to fall into two independent branches (clades a and b) in the ML tree (Fig. 1c). Although the clades a and b remained clustered together in the NJ tree, the bootstrap value lowered to 46.8% compared to the 60.5% in our previous analysis [Liu et al., 1998]. The robustness of the tree topology can be questioned since bootstrap values lower than 75% are generally not confident enough to fully support a topology [Zharkikh and Li, 1992]. For clade c, in spite of a bootstrap value of 65.8% in the NJ tree, it obtained 77.3% bootstrap support in the pars tree and a significant branch length ($P<0.01$) in the ML tree. Since it has been supported by at least two methods, clade c can therefore be considered as true.

All ten Congolese isolates clustered within the GBV-C/HGV type 1 group. No type 2 strain has been detected in Congo in the current study or by others [Smith et al., 1997]. The ten Congolese isolates belong to two different clades. Three of these together with four Ghanaian strains formed a well supported cluster (clade b), whereas the clustering of the other seven Congolese isolates was inconsistent in different methods. These seven strains together with a Belgian isolate formed a well supported monophyletic group (clade d) in the NJ tree, while in the pars tree, clade d only obtained a 34.7% bootstrap support. Moreover, two strains (Z11, Z19) were excluded from this branch in the ML tree. These inconsistent results questioned the robustness of the topology of clade d. Recently, Strimmer and von Haeseler [1997] introduced a new method, likelihood-mapping analysis, that can be used as a complementary approach to solve the controversial phylogenies. This was therefore used to test the reliability of clade d. Two sets of likelihood-mapping statistics showed that most of the quartets were mapped in the top corner of the triangle representing well-resolved phylogenies. This meant that Z11 and Z19 conclusively clustered within clade d. Furthermore, although with low bootstrap support, clade d retained its topology in the pars consensus tree, and a significant branch length ($P<0.01$) in the ML tree. Taken together these data indicate that clade d can represent a fourth at yet undescribed clade in the GBV-C/HGV type 1 sequences.

Inconsistent results were obtained by different phylogenetic analysis methods. It is very difficult to decide which method or methods are most reliable since different data sets may favour different algorithms. Different methods are indeed based on different evolutionary assumptions and if these assumptions are violated by the data sets, the algorithm will not perform optimally. For example, the unweighted pair group method with arithmetic averages (UPGMA), a previously widely used method, assumes that the evolutionary rate is the same in all branches. This assumption is almost never true for viruses. In this case, UPGMA will tend to give a wrong tree when evolutionary rates are very different among different branches. The use of statistical methods can be helpful in estimating the reliability of tree topologies, but if the original data is biased for some reasons, a clade may be regarded as statistically significant although it is a wrong one. Conversely, a cluster may be correct even if its bootstrap value is less than 75%. This is because the statistical methods also depend on the phylogeny methods used and the original bias cannot be corrected by this process. It is advisable to use more than one phylogenetic method with each set of data.

In this study, the prevalence, genomic diversity, and phylogenetic relationship of ten GBV-C/HGV isolates from Congo were investigated. A high prevalence of GBV-C/HGV RNA was found in the population studied. All the Congolese isolates are GBV-C/HGV type 1. Four possible clades were identified among the type 1 sequences. The Congolese strains belonged to 2 clades, one of these undescribed. Whether the four clades will represent genuine subtypes of GBV-C/HGV type 1, cannot at present be concluded. Analysis of more strains from different African countries could confirm the observed trend or on the contrary fade the limits between clades. Another possible bias may be due to the very low number of known genotype 3 sequences. At first sight the presence of multiple evolutionary clades in Africa could plead for an ancient presence of that virus on this continent, but further study is necessary to obtain an insight into evolutionary rates of different genotypes of this virus. This is necessary if conclusions are to be drawn on the general evolution of GBV-C/HGV.

ACKNOWLEDGMENTS

We thank L. Croonen, A. Despineux, M. Michiels, M. Reynders, E. Goderniaux, and M. Smaers for their technical assistance and to G. Burtonboy for encouraging this research.

REFERENCES

Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. 1989. Isolation of a cDNA clone derived from blood-borne non-A, non-B viral hepatitis genome. Science 244:359–362.

Cornu C, Jadoul M, Loute G, Goubau P. 1997. Hepatitis G virus infection in haemodialysed patients: epidemiology and clinical relevance. Nephrol Dial Transplant 12:1326–1329.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.

Felsenstein J. 1989. PHYLIP-Phylogeny inference package (Version 3.2). Cladistics 5:164–166.

Fukushi S, Kurihara C, Ishiyama N, Okamura H, Hoshino FB, Oya A, Katayama K. 1996. Nucleotide sequence of the 5′ noncoding region of hepatitis G virus isolated from Japanese patients: comparison with reported isolates. Biochem Biophys Res Commun 226:314–318.

Kowo MP, Goubau P, Ndam EC, Njoya O, Sasaki S, Seghers V, Kesteloot H. 1995. Prevalence of hepatitis C virus and other bloodborne viruses in Pygmies and neighbouring Bantus in southern Cameroon. Trans Roy Soc Trop Med Hyg 89:484–486.

Leary TP, Muerhoff AS, Simons JN, Pilot-Matias TJ, Erker JC, Chalmers ML, Schlauder GG, Dawson GJ, Desai SM, Mushahwar IK. 1996. Sequence and genomic organisation of GBV-C: a novel member of the Flaviviridae associated with human non-A-E hepatitis. J Med Virol 48:60–67.

Linnen J, Wages J Jr, Zhang-Keck ZY, Fry K, Krawczynski K, Alter H, Koonin E, Gallagher M, Alter M, Hadziyannis S, Karayiannis P, Fung K, Nakatsuji Y, Shih JWK, Young L, Piatak M Jr, Hoover C, Fernandez J, Chen S, Zou JC, Morris T, Hyams KC, Ismay S, Lifson JD, Hess G, Foung SKH, Thomas H, Bradley D, Margolis H, Kim JP. 1996. Molecular cloning and disease association of hepatitis G virus: a transfusion-transmissible agent. Science 271:505–508.

Liu H-F, Cornu C, Jadoul M, Dahan K, Loute G, Goubau P. 1998. Molecular analysis of GB virus C isolates in Belgian haemodialysis patients. J Med Virol 55:118–122.

Masuko K, Mitsui T, Iwano K, Yamazaki C, Okuda K, Meguro T, Murayama N, Inoue T, Tsuda F, Okamoto H, Miyakawa Y, Mayumi M. 1996. Infection with hepatitis GB virus C in patients on maintenance hemodialysis. N Eng J Med 334:1485–1490.

Muerhoff AS, Simons JN, Leary TP, Erker JC, Chalmers ML, Pilot-Matias TJ, Dawson GJ, Desai SM, Mushahwar IK. 1996. Sequence heterogeneity within the 5′-terminal region of the hepatitis GB virus C genome and evidence for genotypes. J Hepatol 25:379–384.

Muerhoff AS, Smith DB, Leary TP, Erker JC, Desai SM, Mushahwar IK. 1997. Identification of GB virus C variants by phylogenetic analysis of 5′-untranslated and coding region sequences. J Virol 71:6501–6508.

Naito H, Win KM, Abe K. 1999. Identification of a novel genotype of hepatitis G virus in southeast Asia. J Clin Microbiol 37:1217–1220.

Saito I, Miyamura T, Ohbayashi A, Harada H, Katayama T, Kikuchi S, Watanabe Y, Koi S, Onji M, Ohta Y, Choo QL, Houghton M, Kuo G. 1990. Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. Proc Natl Acad Sci USA 87: 6547–6549.

Simons JN, Leary TP, Dawson GJ, Pilot-Matias TJ, Muerhoff AS, Schlauder GG, Desai SM, Mushahwar IK. 1995. Isolation of novel virus-like sequences associated with human hepatitis. Nature Med 1:564–569.

Smith DB, Cuceanu N, Davidson F, Jarvis LM, Mokili JLK, Hamid S, Ludlam CA, Simmonds P. 1997. Discrimination of hepatitis G virus/GBV-C geographical variants by analysis of the 5′ non-coding region. J Gen Virol 78:1533–1542.

Strimmer K, von Haeseler A. 1996. Quartet puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969.

Strimmer K, von Haeseler A. 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc Natl Acad Sci USA 94:6815–6819.

Tibbs CJ, Palmer SJ, Coker R, Clark SK, Parsons GM, Hojvat S, Peterson D, Banatvala JE. 1991. Prevalence of hepatitis C in tropical communities: the importance of confirmatory assays. J Med Virology 34:143–147.

Tucker TJ, Louw SJ, Robson SC, Isaacs S, Kirsch RE. 1997. High prevalence of GBV-C hepatitis G virus infection in a rural South African population. J Med Virol 53 225–228.

Zharkikh A, Li W-H. 1992. Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock. Mol Biol Evol 9:1119–1147.

Zuckerman AJ. 1996. Alphabet of hepatitis viruses. Lancet 347:558–559.